Chapter 7

Explainable Large Language Models & iContracts

The Chapter addresses RQ6, which reads as follows:

> **RQ6:** *To what extent is it possible to accelerate the adoption of Intelligent Contracts with Explainable Large Language Models?*

Contract automation is a field of LegalTech under AI and Law that is currently undergoing a transition from Smart to iContracts. iContracts aim toward full contracting automation. Their main challenge is finding a convincing direction for market adoption. Two powerful market factors are the advent of LLMs and AI Regulation. The Chapter investigates how the two factors are able to influence the market adoption of iContracts. Following a literature review our research employs three methodologies: (1) market gap analysis, (2) case study, and (3) application. The results show a clear way for iContracts to follow, based on existing market gaps. Moreover, they validate whether the application of Explainable LLMs is possible. The discussion clarifies the main limitations with Explainable LLMs. Our chapter conclusion is that the two factors are impactful for so long as the market adoption attempts to bridge the gap between innovators and early adopters.

The current chapter corresponds to the following publication:

Stathis, G. (2024b). Explainable Large Language Models & iContracts. *In the Proceedings of the 16th International Conference on Agents and Artificial Intelligence (ICAART)*, 3:1378–1385

7.1 Taking iContracts to the Market

Smart Contracts have laid the foundation for iContracts. Smart Contracts are self-executing contracts with the terms of the agreement directly written into code [Madir, 2020]. They are rooted in blockchain technology, which provides *transparency* and *trust* in the digital realm [Werbach, 2018].

While Smart Contracts have revolutionised the contracting process by automating transactions and reducing the need for intermediaries, they have limitations. Smart Contracts are essentially binary, and capable of executing predefined actions, but incapable of interpreting and adapting to complex legal nuances [Mik, 2017].

iContracts represent the next step in the evolutionary process. They are designed to go beyond the straightforward nature of Smart Contracts. The key objective is to enable iContracts to handle autonomously the entire contracting process, encompassing everything from negotiation to execution. Such an action path entails a seamless integration of human-readable language and code [Mason, 2017]. As a result, iContracts will possess the capacity to understand, adapt, and evolve in response to the intricacies of real-world contracts, thus paving the way for a new era of automation [Stathis et al., 2023d, Stathis et al., 2024].

Meanwhile, the main challenge with iContracts is the market adoption rate [McNamara and Sepasgozar, 2018]. Automation is closely related with the law. Owing to the high legal consequences of iContracts, human users prefer traditional methods such as the direct inclusion of a legal expert [Stathis et al., 2023c]. Such a preference is closely connected with two developing trends: (1) the advent of LLMs and (2) the regulation of AI.

The adoption of LLMs has happened at the speed of light. As a result we may wonder how long the adoption of iContracts will take. Is it five years or only half a year? Next to this question we are facing (1) the global battle to regulate technology [Bradford, 2023] and (2) the consequences of the official positive vote of the European Parliament on AI-Act (July 3, 2024) [1].

One of the main challenges with AI technologies is the lack of user *trustworthiness* [Liang et al., 2022]. Research shows that by understanding *how* AI takes decisions, the outcome of the decision can then be explained, and hence the user trustworthiness will increase [Xu et al., 2019]. Although research also shows that the larger the degree of explainability, the lower the understanding for end users [Ribes et al., 2021].

Our motivation for performing this research is to improve the rate of market adoption of iContracts. We believe that by examining to what extent it is

[1] https://www.walkermorris.co.uk/comment-opinion/landmark-eu-ai-act-finalised/

possible to apply Explainable LLMs on iContracts, we are able to increase user trustworthiness among humans.

The aforementioned thoughts lead us to RQ6.

> **RQ6:** *To what extent is it possible to accelerate the adoption of Intelligent Contracts with Explainable Large Language Models?*

The Chapter's contribution is that (1) it shows how and (2) validates whether the application of Explainable LLMs on iContracts is possible to direct the future attempts in developing iContracts fit for market adoption.

To answer the RQ, we structured the Chapter as follows. In Section 7.2, we describe the relevant literature. Section 7.3 presents the chapter methodology and Section 7.4 provides the results. Section 7.5 discusses these results and focusses on theoretical and practical parameters. Finally, Section 7.6 answers RQ6 and provides the chapter conclusion together with a preview on the future.

7.2 Relevant Literature

The literature present sources on: three revolutions by the development of iContracts (7.2.1), the market adoption in business studies (7.2.2), the application of LLMs in contracts (7.2.3), and the development of trustworthy and Explainable LLMs (7.2.4).

7.2.1 Three Revolutions

The path from human contracts to iContracts is characterised by three revolutions. The *first* revolution started with the transition from physical contracts to digital contracts (also known as Electronic Contracts or eContracts) [Krishna and Karlapalem, 2008]. With the advance of big data, people gradually digitised physical contracts into electronically accessible documents [von Westphalen, 2017]. Moreover, they replaced some physical labor such as signing, by electronic handling. Today, most market developments are concentrating on further adopting and expanding the adoption of digital contracts [2].

In the last decade (2010-2020), with the advent of blockchain technology the *second* revolution occurred. It did switch the focus from digital contracts to smart contracts. Smart contracts are agreements that are executable by code, most often on blockchain-based distributed ledgers [Khan et al., 2021]. The promise of smart contracts is to be seen as the replacement of human language by human code. However, the adoption of smart contracts, beyond specialised

[2] https://www.gartner.com/en/documents/3981321

blockchain or experimental market circles, has not managed to reach an automatically wide market adoption [3]. One of the main challenges of the smart contracts is that it is hard for users to *understand* and *trust* the computer code behind them [Zheng et al., 2020].

Here we arrive at the *third* revolution, the transition from smart contract to iContracts. iContracts aim at *full contracting* automation with minimal to no human involvement [Stathis et al., 2023d, Stathis et al., 2024]. iContracts promise to bridge the gap that smart contracts have been unable to fill. "Bridging" will gain user trustworthiness by combining computer code with user-friendly, readable and understandable code, that originates from physical contracts [McNamara and Sepasgozar, 2020]. Despite the surge in scientific interest in iContracts for the construction industry, there remains a significant demand for traditional contracting (first recognised in 2020 [McNamara, 2020]); meaning the development of iContracts is still in its formative stages.

As a follow-up, we mention a research proposal [Stathis et al., 2023d, Stathis et al., 2024] presented at International Conference on Agents and Artificial Intelligence (ICAART) 2023, which revived research on iContracts by re-purposing use cases, instead of delegating the task of the complex construction industry. A reference to a straightforward freelancer agreement would simplify the evident complexity in construction contracts. In particular, such freelance agreements stipulated the extent to which iContracts may contribute in specific domains. With increasing attention to the details of iContracts in simpler domains it is possible to gradually expand into more complex domains. Notwithstanding these developments, it is still unclear how to develop a path towards the market adoption, that contributes towards end-users increasingly adopting iContracts.

7.2.2 Market Adoption in Business Studies

In business studies, market adoption has been studied for many years [Posthumus et al., 2012, Botha, 2019]. The most important contribution is Geoffrey Moore's *Crossing the Chasm* [Moore and McKenna, 1999] [4]. It studies by meticulously observing (i.e., see through the lenses even with small progress) *how* technologies penetrate a group ranging from (1) innovators to (2) early adopters and gradually advance into (3) wide adopters, (4) late adopters until they reach the so called (5) laggers. The results are general: every technology that enters market adoption follows the same route [Goldasteh et al., 2022]. We are going

[3] https://www.grandviewresearch.com/industry-analysis/smart-contracts-market-report

[4] Here we refer to Geoffrey Moore's Chrossing the Chasm theory and not Gordon Moore's Law, which is the observation that the number of transistors in an integrated circuit double in average every two years.

to examine the adoption of iContracts through the same five lenses (see above). Each contracting revolution can be located at a specific stage in Moore's *Crossing the Chasm*. By identifying the specific position of iContracts relative to digital and smart contracts, it is also possible (1) to identify relevant gaps that must be bridged, and then (2) to investigate the extent to which LLMs can contribute in bridging the gaps.

7.2.3 Large Language Models in Contract Automation

Since the dissemination of decoders with ChatGPT, scientists and business people have attempted to apply decoders (falling under the category of Generative AI) on contracts (see for example recent eighty million fundraising for Silicon Valley based Harvey [5]). The extent and results of such applications are still largely unclear. Here we remark that some research results indicate that decoders perform better than human experts in contract review [Martin et al., 2024]. On the one hand, end-users expect that generative AI tools assist them in developing contracts at a fraction of the costs that it would usually cost to pay a lawyer [6]. On the other hand, the generated contracts (seen as decoders) are not necessarily perceived as *trustworthy*, as with most outputs of decoders and Generative AI [Lenat and Marcus, 2023][7].

They may find results *vague*, *confusing* or even *mysterious* (potentially the result of hallucination). Hence, due to the high impact of legal consequences, end-users (even if they do attempt themselves to generate a contract) still prefer to consult their lawyer [Stathis et al., 2023c]. The *larger* the degree of the severity of legal consequences, the *lower* the reliance on generative contracts [Stathis et al., 2023c]. Some scientists might argue, that in a way Generative Contracts are announcing the *fourth* contracting revolution [Williams, 2024]. From our perspective, we consider Generative Contracts to be only a *feature* (i.e., one aspect) of iContracts (or of Smart and Digital contracts) *and not* the architecture behind iContracts. In the methodology, the result and the discussion sections we will clarify why we believe that this is so.

7.2.4 Trustworthy and Explainable Large Language Models

In 2024, we may expect that the European Union (EU) will approve the very first *rules* for AI in the world [European-Parliament, 2023]. The rules will be presented in the AI Act, a regulation which is aiming to regulate two kinds of

[5] https://siliconangle.com/2023/12/20/harvey-raises-80m-build-generative-ai-legal-professionals/

[6] https://www.docusign.com/blog/products/generative-ai-contracts-agreements

[7] https://www.law360.com/pulse/articles/1789792

AI: Predictive AI and Generative AI [Stathis and van den Herik, 2024]. The main focus of these rules is to make sure AI can be trusted in accordance with the directions provided to the EU by the High-Level Expert Group on AI [High-Level Expert Group on AI, 2019]. Here, we see two sides: (1) society seems to mistrust AI, and (2) the EU wants to protect its people by developing rules to prevent negative consequences by the wide adoption of AI [Lockey et al., 2021]. This is precisely the reason why *some* AI technologies are completely prohibited and *other* AI technologies are considered as High-Risk, which are also subject to most of the regulatory restrictions [8]. *Generative AI* are treated as an exceptional AI technology which should adhere to specific restrictions, due to the high reliance on trained data which are potentially subject to copyright laws [9] [Helberger and Diakopoulos, 2023].

Due to the difficulty in distinguishing levels of *perceived* trustworthiness from *actual* trustworthiness, scientists emphasise the concepts of transparency and accountability in matters related to trustworthy AI [Munn, 2023]. The main principal way along which an AI-system can gain transparency and accountability is *via explainability* [Holzinger et al., 2020]. Explainable AI is the field in which researchers investigate how AI system decisions can be explained in an *understandable*, *interpretable* and *trustworthy* manner for humans [10]. As a result, researchers develop frameworks nad methods to tackle the issue of intepretation, explanation and trustworthiness. Two frameworks developed such purposes are SHapley Additive exPlanations (SHAP), a framework for interpreting predictions of machine learning models [Salih et al., 2024], and Local Interpretable Model Agnostic Explanation (LIME) a technique that explains the predictions of any classifier in interpretable manner [Salih et al., 2024].

In the Springer AI and Ethics Journal Stathis and van den Herik deal with the question: how the explainability issue can be handled with the development of a "culture of explainability" that supports the explicit explanation of the architecture, design, production, and implementation of decision-making across the AI development and implementation chain [Stathis and van den Herik, 2024]. The result is vital for our research, provided that for the case of applying Generative AI on iContracts, such an explainability culture will impact the trustworthiness of the end user significantly. Positive implications include developments for ethical and legal transparency and accountability.

[8] https://www.consilium.europa.eu/en/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/

[9] https://www.consilium.europa.eu/en/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/

[10] https://www.marktechpost.com/2023/03/11/understanding-explainable-ai-and-interpretable-ai/

Last but not least, it should be noted that any attempt towards explanation should take careful consideration the psychological state of end-users [Hoffman et al., 2023]. Explanations may vary depending socio-cultural, psychological and other factors [Belinkov et al., 2020]. Hence, the task of explanation presents more complex challenges than just the seemingly simple and straightforward technical explanation of technical concepts to a technical audience. Users of technology come from diverse backgrounds with different levels of knowledge and understanding, introducing an added layer of complexity for our task [Danilevsky et al., 2020].

7.3 Chapter Methodology

The research methodology starts with an analysis of the market gap (7.3.1) and then introduces the case study (7.3.2) as well the Generative AI application in close coherence with LLMs (7.3.3).

7.3.1 Market Gap Analysis

Our methodology begins with the visualisation of the three revolutions into a graph based on Moore's *Crossing the Chasm* graph (see Figure 7.1). This visualisation will allow us to identify the positioning of the three technologies (*digital*, *smart*, and *intelligent* contracts) in relation to market adoption. Thereafter, we are able to conduct a gap analysis to identify the steps that appear to be the missing links for the adoption of iContracts. Our gap analysis will focus on a specific case study.

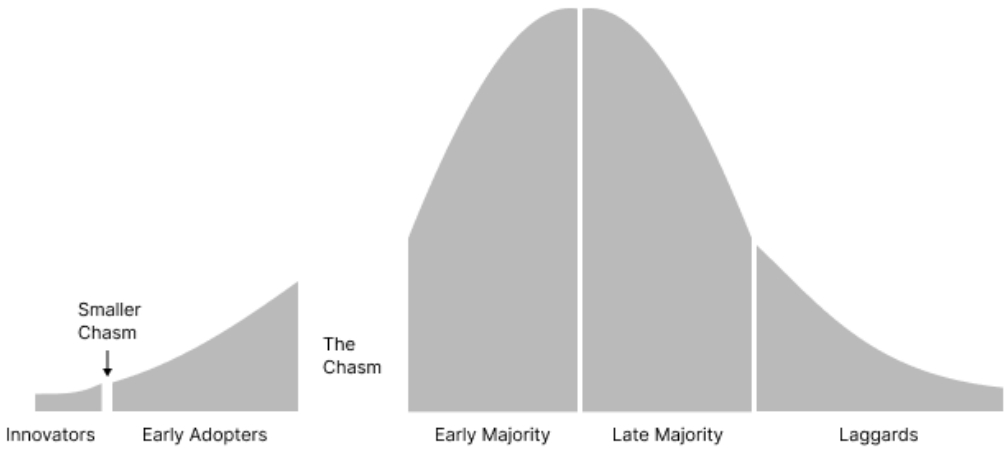

Figure 7.1: Moore's Crossing the Chasm

Figure 7.1 highlights two aspects of the relevant graphs namely (1) Innovators and (2) the Early Adopters by (a) Smaller Chasm and (b) The Chasm. The gaps between Early Majority, Late Majority and Laggards are not examined in detail, because they assume the main chasm has been crossed.

To classify market categories on contracting automation we are going to use Legalcomplex's categorisation [11]. Legalcomplex is the largest database on LegalTech solutions. Legalcomplex has classified contract automation solutions into the following five categories: (1) contract negotiation, (2) contract risk management, (3) contract drafting, (4) contract extraction and (5) contract management.

The classification of Legalcomplex follows the typical journey of a contracting user. Starting with negotiations, a legal expert makes an estimation of legal risk and drafts a contract. Thereafter, relevant elements of that contract can be extracted during the execution stage and/or the monitoring stage; finally the contract is discussed on a management level until its completion.

7.3.2 Case Study

The case study is presented by the Knowledge Graph developed in the ICAART 2023 research and is based on the Onassis Ontology (see Figure 2.1 and Figure 2.2). The Knowledge Graph validates how a small agreement (between two contractors, guided or supervised by a legal expert) can be automated based on *communications* and *risk* data that are exchanged between two contracting parties (Figure 2.2). In the communication with the contractor (the back-end), the role of a legal expert is to clarify the *scope*, *contract*, *risks* and *questions* for the contracting parties. When (1) contracting parties (2) select a scope, (3) the reply questions, and (4) the risk-intelligent contract are updated by the variables from the conversation exchange (see in Figure 2.2 the lines: question, answer, section, contract, variable). The question then is, *how can Generative AI help?* Our experiment will guide the reader to an answer to this question.

7.3.3 Application of a Small Agreement

The application shows to what extent it is possible to develop alternative options by Explainable LLM to facilitate the automated generation of data for the stakeholders involved in our case study (i.e., the legal expert and the contractors). Starting with the KG, we will first identify the potential locations where Generative AI and decoder LLMs can be applied. Then we are showing the power of Google's Gemini LLM to *validate* whether and to what extent it is indeed possible to develop explainable generated data that functionally assist the

[11] https://legalcomplex.com

iContract users into completing their work. To amplify the explainability capabilities of Generative AI and decoder LLMs, in compliance with the literature on trustworthy AI, we will use the opportunity to request the decoder model to provide higher levels of user trustworthiness with alternative explanations.

7.4 Research Results

The results visualise the market gap analysis (7.4.1) and provide the experiment validation (7.4.2).

7.4.1 Market Gap

We start with the market adoption of contracting technologies in accordance with Moore's Crossing the Chasm (Figure 7.1) [Moore and McKenna, 1999]. First, based on the fundamental publication by Moore and McKenna, we find that *physical contracts* are the preferred contracting method for *laggards* (see 7.1 right side). Second, we may understand that *digital contracts* are reaching the *late* market. Here, we remark that *smart contracts* have not managed to achieve a wide adoption of crossing the chasm [Ameyaw et al., 2023]. Third, at the same time, *iContracts* have still *not* reached early adopters [McNamara and Sepasgozar, 2020]. So, at this moment (2024) we can solely state that iContracts are evaluated by innovators only [Stathis et al., 2023d].

Based on (1) this visualisation and (2) the user preferences, we are able to identify the relevant market gap among each contracting alternative, on the basis of five categories: (1) negotiation, (2) risk management, (3) drafting, (4) extraction, and (5) management (Table 7.1 Column 1). Using as comparison framework Legalcomplex's classification of contract automation solutions, we see in Table 7.1 (column 2) that the *highest* adoption is observed with *physical* contracts. The main obstacle for physical contracts is *data extraction*. That obstacle has been solved by *digital* contracts. Hence, gradually digital contracts are the *first followers*. Still digital contracts have *not* managed to replace negotiations or risk management as they occur in physical contracts. Also contract drafting seems to have remained significantly reliant on physical contract drafting (although it is nowadays done often by electronic means, even with the use of templates). *Smart* contracts (see column 4) have two lower levels of adoption across the categories. These are developments to be noted with drafting, extraction and management-which is obviously amplified by electronic means. With *intelligent* contracts, we see the lower adoption rate. The Table 7.1 helps us to find which gap iContracts should aim to bridge first. That is the gap in contract negotiations, risk management and drafting. As seen in Figure 2.2, this is the conceptual line of execution that the Onassis Ontology is taking. Hence,

by experimenting with the application of Explainable Generative AI and LLMs on the Onassis Ontology, it is expected to bridge the market gap observed on these three categories faster.

Table 7.1: Rate of Contract Technology Adoption

	Physical	**Digital**	**Smart**	**Intelligent**
Contract Negotiation	High	Mid	Low	Low
Contract Risk Management	High	Mid	Low	Low
Contract Drafting	High	Mid	Mid	Low
Contract Extraction	Low	High	Mid	Low
Contract Management	High	High	Mid	Low

7.4.2 Explainable Generative AI Validation

The application results, which are accessible on GitHub ([12]) validate that Generative AI can be leveraged in every single step of the Onassis Ontology. For as long as the Generative AI prompt requires the specification of supporting explanations, the provision of such explanations is possible. As a direct consequence, we see that it becomes visible that the *reduction of human labour* in iContracts is further amplified with the help of Generative AI. By including the explicit provision of explanations we may also expect that the user trustworthiness increases. Hence, technologically, it is possible to support the gap difference between iContracts with other contracting alternatives (see Table 7.1) with Explainable Generative AI.

7.5 Discussion

In our discussion we focus on the implications for the adoption of iContracts (7.5.1) and the limitations observed in the Explainable LLM application (7.5.2).

7.5.1 iContracts Market Adoption Implications

As we have seen in Subsection 7.3.1, the market gap analysis became useful in identifying the precise gap between the alternative contracting options for end users. We saw that the advantage of iContracts lies exactly where *digital* and *smart* contracts seem to struggle with replacing physical contracts, namely contract (1) negotiations, (2) risk management and (3) drafting.

[12] https://github.com/onassisontology/onassisontology/\blob/main/img/EGENAIEXP.png

In compliance with the market adoption, the business-study literature showed that physical, digital and smart contracts already present certain alternatives that users have habituated. The opportunity for iContracts to have an impact on early adopters and gradually wide adopters lies in exploiting opportunities that significantly impact current practices.

In conclusion, there might be an opportunity by applying Explainable Generative AI, which simplifies human complexity in multiple directions. This has been validated by its application on the case study. Still, such application allows us also to identify specific limitations which we present below.

7.5.2 Explainable Large Language Model Limitations

Requesting a decoder to develop supporting explanations is an achievable challenge. Such explanations improve user trustworthiness by supporting their evaluation process. Instead of requesting a decoder to directly provide a single answer for a given topic, the alternative option is to allow for users to select the most serviceable answer for a specific purpose.

If we compare the effectivity of (1) humans to produce alternative explanations with (2) a decoder producing such alternatives, we find that it is a laborious task that most humans would avoid due to increased complexity. An alternative conclusion might be that an Explainable LLM is able to replace a laborious task with a *more trustworthy* alternative.

An additional positive aspect is that we validated the application of Explainable LLM as being possible across each step of the Onassis Ontology. Indeed, we found by the complexity of the contracting process and the relatively low impact degree of decoder that the application of Generative AI on iContracts happens on *feature level* and *not on architecture level*. All in all, our request for explanations still presents certain limitations, which are not yet solved.

The limitations are related to the inability of Generative AI systems to provide (1) readable sources that verify its explanations and (2) understandable reasoning patterns that explain the reasoning of its developers. In compliance with the "culture of explanation" as supported in [Stathis and van den Herik, 2024], these two characteristics will be able to drastically increase the user trustworthiness even further, due to the ability for users to validate the origin of specific data as well as the line of reasoning a developer has followed. Such transparency helps with (1) improving the end-user's evaluations and (2) assigning liability with improved accountability that also increases user trustworthiness from a legal point of view.

7.6 Chapter Conclusion

The RQ6, addressed in this Chapter, is:

> **RQ6:** *To what extent is it possible to accelerate the adoption of Intelligent Contracts with Explainable Large Language Models?*

The answer is that Explainable LLMs, as a category of Generative AI, have the possibility to accelerate the adoption of iContracts, namely contract (1) negotiations, (2) risk management and (3) drafting. This is so, because of two characteristics for Explainable LLMs. *First*, Generative AI can make the laborious human tasks involved in the three categories simpler. *Second*, the explainability aspect can increase the end-user trustworthiness significantly by supporting the outputs of Generative AI with explicit explanations. To further leverage explainability, we advise compliance with the "culture of explanations". Hence, the developers of decoders should be *connoisseurs* of the culture in which the system operates when given the task to make an explicit presentation of sources supporting data outputs as well as explicit representation of the line of reasoning supporting algorithmic models.

Our further research will focus on developing experiments with additional case studies, with gradually increasing complexity. Owing to a high reliance on end user reactions, we will continue to conduct end user *validation experiments with graphical user interfaces*.

The overall novelties of the Chapter are (1) presenting the state of *market adoption* of contracting technologies, (2) identifying the *gap among the alternative* contracting technologies, (3) connecting the aspect of *explainability with decoders*, and (4) concluding that the application of Explainable LLMs is successfully possible on practical case studies.